

0590
0620

8



OIPE

ENTERED

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/977,261

DATE: 06/18/2002 P.6
TIME: 09:12:20

Input Set : A:\38621259.app

Output Set: N:\CRF3\06182002\I977261.raw

```

3 <110> APPLICANT: ULLRICH, AXEL
4     GISHIZKY, MIKHAIL
5     SURES, IRMINGARD
7 <120> TITLE OF INVENTION: NOVEL MEGAKARYOCYTIC PROTEIN TYROSINE KINASES
9 <130> FILE REFERENCE: 038602/1259
11 <140> CURRENT APPLICATION NUMBER: 09/977,261
12 <141> CURRENT FILING DATE: 2001-10-16
14 <150> PRIOR APPLICATION NUMBER: 08/232,545
15 <151> PRIOR FILING DATE: 1994-04-22
17 <160> NUMBER OF SEQ ID NOS: 24
19 <170> SOFTWARE: PatentIn Ver. 2.1
21 <210> SEQ ID NO: 1
22 <211> LENGTH: 2000
23 <212> TYPE: DNA
24 <213> ORGANISM: Unknown Organism
26 <220> FEATURE:
27 <221> NAME/KEY: CDS
28 <222> LOCATION: (258)..(1778)
30 <220> FEATURE:
31 <223> OTHER INFORMATION: Description of Unknown Organism: Megakaryocyte
32     kinase 1
34 <400> SEQUENCE: 1
35 ctcgctccaa gttgtgcagc cgggaccgcc tgggggtgtg cagccggctc gcggaggccc 60
37 tcctgggggc gggcgcgggg cggctcgggg gcgccccctg agcagaaaac aggaagaacc 120
39 aggtctcggtc cagtggcacc cagctcccta cctcctgtgc cagccgctg gcctgtggca 180
41 ggccattccc agcgtccccg actgtgacca cttgtcagt gtgcctctca cctgcctcag 240
43 tttccctctg gggggcgt atg gcg ggg cga ggc tct ctg gtt tcc tgg cgg 290
44     Met Ala Gly Arg Gly Ser Leu Val Ser Trp Arg
45     1 5 10
47 gca ttt cac ggc tgt gat tct gct gag gaa ctt ccc cgg gtg agc ccc 338
48 Ala Phe His Gly Cys Asp Ser Ala Glu Glu Leu Pro Arg Val Ser Pro
49     15 20 25
51 cgc ttc ctc cga gcc tgg cac ccc cct ccc gtc tca gcc agg atg cca 386
52 Arg Phe Leu Arg Ala Trp His Pro Pro Pro Val Ser Ala Arg Met Pro
53     30 35 40
55 acg agg cgc tgg gcc ccg ggc acc cag tgt atc acc aaa tgc gag cac 434
56 Thr Arg Arg Trp Ala Pro Gly Thr Gln Cys Ile Thr Lys Cys Glu His
57     45 50 55
59 acc cgc ccc aag cca ggg gag ctg gcc ttc cgc aag ggc gac gtg gtc 482
60 Thr Arg Pro Lys Pro Gly Glu Leu Ala Phe Arg Lys Gly Asp Val Val
61 60 65 70 75
63 acc atc ctg gag gcc tgc gag aac aag agc tgg tac cgc gtc aag cac 530
64 Thr Ile Leu Glu Ala Cys Glu Asn Lys Ser Trp Tyr Arg Val Lys His

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/977,261

DATE: 06/18/2002

TIME: 09:12:20

Input Set : A:\38621259.app

Output Set: N:\CRF3\06182002\I977261.raw

65		80		85		90	
67	cac acc agt gga cag gag ggg ctg ctg gca gct ggg gcg ctg cgg gag						578
68	His Thr Ser Gly Gln Glu Gly Leu Leu Ala Ala Gly Ala Leu Arg Glu						
69		95		100		105	
71	cgg gag gcc ctc tcc gca gac ccc aag ctc agc ctc atg ccg tgg ttc						626
72	Arg Glu Ala Leu Ser Ala Asp Pro Lys Leu Ser Leu Met Pro Trp Phe						
73		110		115		120	
75	cac ggg aag atc tcg ggc cag gag gct gtc cag cag ctg cag cct ccc						674
76	His Gly Lys Ile Ser Gly Gln Glu Ala Val Gln Gln Leu Gln Pro Pro						
77		125		130		135	
79	gag gat ggg ctg ttc ctg gtg cgg gag tcc gcg cgc cac ccc ggc gac						722
80	Glu Asp Gly Leu Phe Leu Val Arg Glu Ser Ala Arg His Pro Gly Asp						
81	140		145		150		155
83	tac gtc ctg tgc gtg agc ttt ggc cgc gac gtc atc cac tac cgc gtg						770
84	Tyr Val Leu Cys Val Ser Phe Gly Arg Asp Val Ile His Tyr Arg Val						
85		160		165		170	
87	ctg cac cgc gac ggc cac ctc aca atc gat gag gcc gtg ttc ttc tgc						818
88	Leu His Arg Asp Gly His Leu Thr Ile Asp Glu Ala Val Phe Phe Cys						
89		175		180		185	
91	aac ctc atg gac atg gtg gag cat tac agc aag gac aag ggc gct atc						866
92	Asn Leu Met Asp Met Val Glu His Tyr Ser Lys Asp Lys Gly Ala Ile						
93		190		195		200	
95	tgc acc aag ctg gtg aga cca aag cgg aaa cac ggg acc aag tcg gcc						914
96	Cys Thr Lys Leu Val Arg Pro Lys Arg Lys His Gly Thr Lys Ser Ala						
97		205		210		215	
99	gag gag gag ctg gcc agg gcg ggc tgg tta ctg aac ctg cag cat ttg						962
100	Glu Glu Glu Leu Ala Arg Ala Gly Trp Leu Leu Asn Leu Gln His Leu						
101	220		225		230		235
103	aca ttg gga gca cag atc gga gag gga gag ttt gga gct gtc ctg cag						1010
104	Thr Leu Gly Ala Gln Ile Gly Glu Gly Glu Phe Gly Ala Val Leu Gln						
105		240		245		250	
107	ggt gag tac ctg ggg caa aag gtg gcc gtg aag aat atc aag tgt gat						1058
108	Gly Glu Tyr Leu Gly Gln Lys Val Ala Val Lys Asn Ile Lys Cys Asp						
109		255		260		265	
111	gtg aca gcc cag gcc ttc ctg gac gag acg gcc gtc atg acg aag atg						1106
112	Val Thr Ala Gln Ala Phe Leu Asp Glu Thr Ala Val Met Thr Lys Met						
113		270		275		280	
115	caa cac gag aac ctg gtg cgt ctc ctg ggc gtg atc ctg cac cag ggg						1154
116	Gln His Glu Asn Leu Val Arg Leu Leu Gly Val Ile Leu His Gln Gly						
117		285		290		295	
119	ctg tac att gtc atg gag cac gtg agc aag ggc aac ctg gtg aac ttt						1202
120	Leu Tyr Ile Val Met Glu His Val Ser Lys Gly Asn Leu Val Asn Phe						
121	300		305		310		315
123	ctg cgg acc cgg ggt cga gcc ctc gtg aac acc gct cag ctc ctg cag						1250
124	Leu Arg Thr Arg Gly Arg Ala Leu Val Asn Thr Ala Gln Leu Leu Gln						
125		320		325		330	
127	ttt tct ctg cac gtg gcc gag ggc atg gag tac ctg gag agc aag aag						1298
128	Phe Ser Leu His Val Ala Glu Gly Met Glu Tyr Leu Glu Ser Lys Lys						
129		335		340		345	

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/977,261

DATE: 06/18/2002

TIME: 09:12:20

Input Set : A:\38621259.app

Output Set: N:\CRF3\06182002\I977261.raw

```

131 ctt gtg cac cgc gac ctg gcc gcc cgc aac atc ctg gtc tca gag gac 1346
132 Leu Val His Arg Asp Leu Ala Ala Arg Asn Ile Leu Val Ser Glu Asp
133      350      355      360
135 ctg gtg gcc aag gtc agc gac ttt ggc ctg gcc aaa gcc gag cgg aag 1394
136 Leu Val Ala Lys Val Ser Asp Phe Gly Leu Ala Lys Ala Glu Arg Lys
137      365      370      375
139 ggg cta gac tca agc cgg ctg ccc gtc aag tgg acg gcg ccc gag gct 1442
140 Gly Leu Asp Ser Ser Arg Leu Pro Val Lys Trp Thr Ala Pro Glu Ala
141 380      385      390      395
143 ctc aaa cac ggg aag ttc acc agc aag tcg gat gtc tgg agt ttt ggg 1490
144 Leu Lys His Gly Lys Phe Thr Ser Lys Ser Asp Val Trp Ser Phe Gly
145      400      405      410
147 gtg ctg ctc tgg gag gtc ttc tca tat gga cgg gct ccg tac cct aaa 1538
148 Val Leu Leu Trp Glu Val Phe Ser Tyr Gly Arg Ala Pro Tyr Pro Lys
149      415      420      425
151 atg tca ctg aaa gag gtg tcg gag gcc gtg gag aag ggg tac cgc atg 1586
152 Met Ser Leu Lys Glu Val Ser Glu Ala Val Glu Lys Gly Tyr Arg Met
153      430      435      440
155 gaa ccc ccc gag ggc tgt cca ggc ccc gtg cac gtc ctc atg agc agc 1634
156 Glu Pro Pro Glu Gly Cys Pro Gly Pro Val His Val Leu Met Ser Ser
157      445      450      455
159 tgc tgg gag gca gag ccc gcc cgc cgg cca ccc ttc cgc aaa ctg gcc 1682
160 Cys Trp Glu Ala Glu Pro Ala Arg Arg Pro Pro Phe Arg Lys Leu Ala
161 460      465      470      475
163 gag aag ctg gcc cgg gag cta cgc agt gca ggt gcc cca gcc tcc gtc 1730
164 Glu Lys Leu Ala Arg Glu Leu Arg Ser Ala Gly Ala Pro Ala Ser Val
165      480      485      490
167 tca ggg cag gac gcc gac ggc tcc acc tcg ccc cga agc cag gag ccc 1778
168 Ser Gly Gln Asp Ala Asp Gly Ser Thr Ser Pro Arg Ser Gln Glu Pro
169      495      500      505
171 tgacccacc cgggtggggcc cttggcccca gaggaccgag agagtggaga gtgcggcgtg 1838
173 ggggcactga ccaggcccaa ggagggtcca ggcgggcaag tcatcctcct ggtgccca 1898
175 gcaggggctg gccacgtag ggggctctgg gcggcccgag gacacccag acctgcgaag 1958
177 gatgatgcc cgataaagac ggattctaag gactctaaaa aa 2000
180 <210> SEQ ID NO: 2
181 <211> LENGTH: 507
182 <212> TYPE: PRT
183 <213> ORGANISM: Unknown Organism
185 <220> FEATURE:
186 <223> OTHER INFORMATION: Description of Unknown Organism: Megakaryocyte
187      kinase 1
189 <400> SEQUENCE: 2
190 Met Ala Gly Arg Gly Ser Leu Val Ser Trp Arg Ala Phe His Gly Cys
191      1      5      10      15
193 Asp Ser Ala Glu Glu Leu Pro Arg Val Ser Pro Arg Phe Leu Arg Ala
194      20      25      30
196 Trp His Pro Pro Pro Val Ser Ala Arg Met Pro Thr Arg Arg Trp Ala
197      35      40      45
199 Pro Gly Thr Gln Cys Ile Thr Lys Cys Glu His Thr Arg Pro Lys Pro

```

RAW SEQUENCE LISTING

DATE: 06/18/2002

PATENT APPLICATION: US/09/977,261

TIME: 09:12:20

Input Set : A:\38621259.app

Output Set: N:\CRF3\06182002\I977261.raw

200	50	55	60													
202	Gly	Glu	Leu	Ala	Phe	Arg	Lys	Gly	Asp	Val	Val	Thr	Ile	Leu	Glu	Ala
203	65					70					75				80	
205	Cys	Glu	Asn	Lys	Ser	Trp	Tyr	Arg	Val	Lys	His	His	Thr	Ser	Gly	Gln
206				85					90						95	
208	Glu	Gly	Leu	Leu	Ala	Ala	Gly	Ala	Leu	Arg	Glu	Arg	Glu	Ala	Leu	Ser
209				100					105					110		
211	Ala	Asp	Pro	Lys	Leu	Ser	Leu	Met	Pro	Trp	Phe	His	Gly	Lys	Ile	Ser
212				115				120					125			
214	Gly	Gln	Glu	Ala	Val	Gln	Gln	Leu	Gln	Pro	Pro	Glu	Asp	Gly	Leu	Phe
215		130					135					140				
217	Leu	Val	Arg	Glu	Ser	Ala	Arg	His	Pro	Gly	Asp	Tyr	Val	Leu	Cys	Val
218	145					150				155					160	
220	Ser	Phe	Gly	Arg	Asp	Val	Ile	His	Tyr	Arg	Val	Leu	His	Arg	Asp	Gly
221				165					170						175	
223	His	Leu	Thr	Ile	Asp	Glu	Ala	Val	Phe	Cys	Asn	Leu	Met	Asp	Met	
224				180					185					190		
226	Val	Glu	His	Tyr	Ser	Lys	Asp	Lys	Gly	Ala	Ile	Cys	Thr	Lys	Leu	Val
227		195						200					205			
229	Arg	Pro	Lys	Arg	Lys	His	Gly	Thr	Lys	Ser	Ala	Glu	Glu	Glu	Leu	Ala
230		210					215					220				
232	Arg	Ala	Gly	Trp	Leu	Leu	Asn	Leu	Gln	His	Leu	Thr	Leu	Gly	Ala	Gln
233	225					230					235				240	
235	Ile	Gly	Glu	Gly	Glu	Phe	Gly	Ala	Val	Leu	Gln	Gly	Glu	Tyr	Leu	Gly
236				245						250					255	
238	Gln	Lys	Val	Ala	Val	Lys	Asn	Ile	Lys	Cys	Asp	Val	Thr	Ala	Gln	Ala
239				260					265					270		
241	Phe	Leu	Asp	Glu	Thr	Ala	Val	Met	Thr	Lys	Met	Gln	His	Glu	Asn	Leu
242				275				280						285		
244	Val	Arg	Leu	Leu	Gly	Val	Ile	Leu	His	Gln	Gly	Leu	Tyr	Ile	Val	Met
245		290					295					300				
247	Glu	His	Val	Ser	Lys	Gly	Asn	Leu	Val	Asn	Phe	Leu	Arg	Thr	Arg	Gly
248	305					310					315				320	
250	Arg	Ala	Leu	Val	Asn	Thr	Ala	Gln	Leu	Leu	Gln	Phe	Ser	Leu	His	Val
251				325						330					335	
253	Ala	Glu	Gly	Met	Glu	Tyr	Leu	Glu	Ser	Lys	Lys	Leu	Val	His	Arg	Asp
254				340					345					350		
256	Leu	Ala	Ala	Arg	Asn	Ile	Leu	Val	Ser	Glu	Asp	Leu	Val	Ala	Lys	Val
257			355				360						365			
259	Ser	Asp	Phe	Gly	Leu	Ala	Lys	Ala	Glu	Arg	Lys	Gly	Leu	Asp	Ser	Ser
260		370					375					380				
262	Arg	Leu	Pro	Val	Lys	Trp	Thr	Ala	Pro	Glu	Ala	Leu	Lys	His	Gly	Lys
263	385					390					395				400	
265	Phe	Thr	Ser	Lys	Ser	Asp	Val	Trp	Ser	Phe	Gly	Val	Leu	Leu	Trp	Glu
266				405						410					415	
268	Val	Phe	Ser	Tyr	Gly	Arg	Ala	Pro	Tyr	Pro	Lys	Met	Ser	Leu	Lys	Glu
269				420					425					430		
271	Val	Ser	Glu	Ala	Val	Glu	Lys	Gly	Tyr	Arg	Met	Glu	Pro	Pro	Glu	Gly
272			435					440						445		

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/977,261

DATE: 06/18/2002

TIME: 09:12:20

Input Set : A:\38621259.app

Output Set: N:\CRF3\06182002\I977261.raw

```

274 Cys Pro Gly Pro Val His Val Leu Met Ser Ser Cys Trp Glu Ala Glu
275      450      455      460
277 Pro Ala Arg Arg Pro Pro Phe Arg Lys Leu Ala Glu Lys Leu Ala Arg
278 465      470      475      480
280 Glu Leu Arg Ser Ala Gly Ala Pro Ala Ser Val Ser Gly Gln Asp Ala
281      485      490      495
283 Asp Gly Ser Thr Ser Pro Arg Ser Gln Glu Pro
284      500      505
287 <210> SEQ ID NO: 3
288 <211> LENGTH: 2500
289 <212> TYPE: DNA
290 <213> ORGANISM: Unknown Organism
292 <220> FEATURE:
293 <221> NAME/KEY: CDS
294 <222> LOCATION: (82)..(2106)
296 <220> FEATURE:
297 <223> OTHER INFORMATION: Description of Unknown Organism: Megakaryocyte
298      kinase 2
300 <400> SEQUENCE: 3
301 ccgcttttttg cttagagctt gagagtcaaa gttaaggacc cacatgtata cttcggtctt 60
303 agcgagtcta aggatgataa t atg gat aca aaa tct att cta gaa gaa ctt 111
304      Met Asp Thr Lys Ser Ile Leu Glu Glu Leu
305      1      5      10
307 ctt ctc aaa aga tca cag caa aag aag aaa atg tca cca aat aat tac 159
308 Leu Leu Lys Arg Ser Gln Gln Lys Lys Lys Met Ser Pro Asn Asn Tyr
309      15      20      25
311 aaa gaa cgg ctt ttt gtt ttg acc aaa aca aac ctt tcc tac tat gaa 207
312 Lys Glu Arg Leu Phe Val Leu Thr Lys Thr Asn Leu Ser Tyr Tyr Glu
313      30      35      40
315 tat gac aaa atg aaa agg ggc agc aga aaa gga tcc att gaa att aag 255
316 Tyr Asp Lys Met Lys Arg Gly Ser Arg Lys Gly Ser Ile Glu Ile Lys
317      45      50      55
319 aaa atc aga tgt gtg gag aaa gta aat ctc gag gag cag acg cct gta 303
320 Lys Ile Arg Cys Val Glu Lys Val Asn Leu Glu Glu Gln Thr Pro Val
321      60      65      70
323 gag aga cag tac cca ttt cag att gtc tat aaa gat ggg ctt ctc tat 351
324 Glu Arg Gln Tyr Pro Phe Gln Ile Val Tyr Lys Asp Gly Leu Leu Tyr
325 75      80      85      90
327 gtc tat gca tca aat gaa gag agc cga agt cag tgg ttg aaa gca tta 399
328 Val Tyr Ala Ser Asn Glu Glu Ser Arg Ser Gln Trp Leu Lys Ala Leu
329      95      100      105
331 caa aaa gag ata agg ggt aac ccc cac ctg ctg gtc aag tac cat agt 447
332 Gln Lys Glu Ile Arg Gly Asn Pro His Leu Leu Val Lys Tyr His Ser
333      110      115      120
335 ggg ttc ttc gtg gac ggg aag ttc ctg tgt tgc cag cag agc tgt aaa 495
336 Gly Phe Phe Val Asp Gly Lys Phe Leu Cys Cys Gln Gln Ser Cys Lys
337      125      130      135
339 gca gcc cca gga tgt acc ctc tgg gaa gca tat gct aat ctg cat act 543
340 Ala Ala Pro Gly Cys Thr Leu Trp Glu Ala Tyr Ala Asn Leu His Thr

```

RAW SEQUENCE LISTING ERROR SUMMARY
PATENT APPLICATION: US/09/977,261

DATE: 06/18/2002
TIME: 09:12:21

Input Set : A:\38621259.app
Output Set: N:\CRF3\06182002\I977261.raw

Please Note:

Use of n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220> to <223> fields of each sequence which presents at least one n or Xaa.

Seq#:20; N Pos. 14,20,23,26

Seq#:21; N Pos. 12,18,24

Seq#:24; Xaa Pos. 6

VERIFICATION SUMMARY

PATENT APPLICATION: US/09/977,261

DATE: 06/18/2002

TIME: 09:12:21

Input Set : A:\38621259.app

Output Set: N:\CRF3\06182002\I977261.raw

L:2362 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:20 after pos.:0

L:2389 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:21 after pos.:0

L:2435 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:24 after pos.:0